Multimed Tools Appl
DOI 10.1007/s11042-008-0216-1

Semantic force relevance feedback, content-free 3D object retrieval and annotation propagation: bridging the gap and beyond

Efstathios Onasoglou · Petros Daras

Abstract Relevance Feedback is a technique used for enhancing retrieval accuracy in multimedia database systems. In this paper two novel relevance feedback algorithms are proposed for 3D object databases, in which the relative scores of various users, which express users' subjectivity, are kept accumulatively as additional descriptors. Each object is interpreted as a charged particle, whose relative scores represent the value of the charge. Based on these charges, semantic forces are calculated between the 3D objects, which are repelled or attracted properly in the feature space. The forces are of dual nature, semantic and geometric, in the first algorithm, whereas they are purely semantic in the second one. Furthermore, a novel algorithm for annotation propagation is developed, which is based on a linear prediction scheme of the changes that must be made in the feature vector of a newly added 3D object in the database, according to alterations that has already taken place to objects in the database. The combination of low and high level features in one formula, is able to fill the semantic gap as much as possible till time being, while the proposed content-free retrieval method illustrates the fact that in the long run, a purely semantic algorithm can provide excellent retrieval results.

Keywords Relevance feedback · Semantic propagation · Clustering

This work was supported by the VICTORY EU IST project, cn. 044985.

E. Onasoglou (✉) · P. Daras
Informatics and Telematics Institute, Centre for Research and Technology Hellas,
1st km Thermi-Panorama Rd, P.O. Box 60361, 57001, Thessaloniki, Greece
e-mail: onaso@iti.gr

P. Daras
e-mail: daras@iti.gr

1 Introduction

The advances in computer technology and more specifically, in multimedia capturing technology have led to storing, transmitting and generally using, massive amounts of multimedia data.

In modern 3D databases, there is an imperative need for accurate retrieval of objects that are semantically similar to user queries, in order to reuse 3D objects and to exploit knowledge relative to them.

The most powerful technique in improving retrieval adequacy, apart from devising better low-level descriptors, is the use of Relevance Feedback (RF). In RF schemes, the user becomes an active part of the retrieval system. Firstly, s/he supplies the system with a query. The system responses with a list of similar objects. The user gives relevance scores to a number of objects of the list according to his/her own judgement. Then, these scores are fed back to the system, which refines the search and responds with a new list that contains objects which are more similar to the user query.

The motivation for developing the first proposed RF method that is presented in this paper, was the fact that there is currently no RF method, to the extend of the authors' knowledge, which combines both low-level and semantic features, in one single formula, in order to improve the retrieval results. This combination leads to improved results as it is proven experimentally in Section 6.1. Furthermore, the problem of pure semantic retrieval is a current research topic that has, still, been minimally explored and is addressed in this paper with an innovative RF method which provides promising results. Moreover, the problem of adding new objects in a multimedia database, which must keep up with the changes made by an RF algorithm is rarely confronted in previous research. Lastly, the main motivation of the presented work was to present a combined high/low-level RF method, a purely semantic one and a method to avoid the cold start problem, all in one framework, something that has not been addressed so far.

The rest of the paper is organized as follows: In Section 2 the state-of-the-art of current RF schemes and the authors' proposed methods are presented. In Section 3 the Semantic Force Relevance Feedback (SFRF) algorithm is presented. In Section 4 the Pure Semantic Relevance Feedback (PSRF) algorithm is given and in Section 5 is described the proposed solution to the cold start problem. Further, in Section 6 are presented the experimental results, which illustrate the performance of the proposed schemes. Finally, conclusions are drawn in Section 7.

2 State-Of-the-art, description of the proposed methods and added value

In image retrieval, the first RF systems were computer centric [10] and the user had to know the interconnection of low- and high-level features, which posed problems as the subjectivity of human perception. In [22], objects were represented by various properties (e.g. colour) and each property was described by feature vectors. Weights were assigned to the properties, according to which draws the user's attention and to different vector dimensions, according to the deviation of each dimension. In [13], the problem of RF was, formulated, also, as feature component weighting, but the distance metric that was used, was the generalized ellipsoid distance:

$D(\vec{x}, \vec{q}) = (\vec{x} - \vec{q})^T M(\vec{x} - \vec{q})$, where $\vec{q}$ is the feature vector of the query, $\vec{x}$ is the corresponding vector of an object and M a symmetric weighting matrix. In [28], was presented an automated method to update the weights used in [22] that utilized a support vector machine (SVM) [31] scheme. SVM techniques were also used for RF in [12, 29], where SVM were used for soft labelling and annotation propagation. Bayesian methods have been used in [8], where the probability of all objects in the database of being relevant were estimated and objects were presented to the user according to the estimated probability. In [27], the assumption that positive examples belong to the same Gaussian class was made. Discriminant functions [26] were estimated and used. Other methods include probabilistic frameworks such as in [7], where the data were annotated with binary vectors indicating presence or absence of a quality and the conditional probability of the query being relevant to an object with the condition of it, being relevant to a known set of objects, was estimated with Bayes rules.

The aforementioned methods are mainly focused in the short term learning of the semantic class of the query. In order to improve the RF performance, relying on the knowledge gained in the long run there have been developed methods that act based on the feedback of all user queries. In [6], images in the database are viewed as “bags” and the regions of interest as instances, something that leads to a combination of the Multiple Instance Learning method and the training of a neural network, whose weights are being updated in every RF step. In [11] and [21], a matrix is held for storing the scores. Relying on this matrix, those methods proceed to weight updating and ranking.

Concerning the case of 3D object retrieval, only few RF methods have been proposed. One of the most noticeable ones is presented in [2], where each object in the database is moved by a certain amount towards or away from the query, depending on how far it is located from positive or negative examples, respectively. In this way the descriptors change and the objects are, permanently, displaced in new positions in the feature space, hopefully creating proper semantic classes. In [5], snapshots of each 3D object are taken in different angles and then image RF techniques are implemented. In [16], the Kernel Principal Component Analysis (KPCA) [24] and the Linear Discriminant Analysis (LDA) [26] algorithms are implemented in order to improve the retrieval accuracy during RF iterations. A different approach is given in [1], where query processing and RF are performed using pre-computed, pair-wise distances between objects. It must be clarified that, although, recently the effort in the RF research has moved to 3D data, the core framework of text or image RF methods is also applicable to the 3D case. The main difference in the retrieval of various kinds of data is the procedure of low-level feature extraction, which is radically different among different kind of multimedia data. After this indexing procedure and the placement of the objects (documents, images, 3D data etc.) in a feature space, any RF method can be used. Of course the parameters of each RF method, must be optimized for each case of multimedia data.

RF methods are, traditionally viewed as supporting algorithms for a main content-based framework. Recently, a few efforts have taken place, which are aiming at the detachment of algorithms from the necessity of descriptor extraction. In [14, 17], are presented two quite similar methods that rely on accumulation of records of user feedback. The system recycles them in the form of collaborative filtering [33]. The collaborative filtering technique helps the system to extract probability values, which describe the grade of inter-object relevance, based on, per user, sets of selected objects. Furthermore, as far as annotation propagation is concerned, few works have been published. In [35], for example, a semantic network is trained in order to propagate keywords to images with the RF procedure. Also, noteworthy is the method in [34], where RF and kernel regression [19] algorithms are used for propagating hidden annotation. The disadvantage of these methods is that either they equate semantics with keywords and thus they deteriorate the role of semantic-driven retrieval since keywords are not a pure semantic property, or they are heavily based onto geometry.

In this paper, two novel RF methods are proposed. The first method is called SFRF and the second PSRF. Both algorithms assume action of the 3D objects similar to charged particles that apply coulomb-like forces to each other. These forces are of such nature that repel the semantically irrelevant objects and attract the relevant ones. The charge of these "particles" is the relevant scores that they have been gained from the users' feedback through all sessions for all users. That way, during SFRF, the initial geometrical object descriptors are changed so that the objects are moved in positions, in the feature space, where semantic classes are more discernible. In the PSRF algorithm there are no geometric descriptors and the attractive or repulsive forces are calculated only by the relevance scores. In PSRF, the vector of relevance scores of each 3D object, is also the object's feature vector and thus the conceptual multi-dimensional feature space, where all objects belong to, is formed only by the users' preference.

SFRF is inherently advantageous to methods that do not keep "history" of feedback sessions because it accumulates and exploits knowledge in the run of all sessions. Furthermore, SFRF is computationally inexpensive and very faster than methods that use SVM or other forms of resource-demanding training ([12, 29]). These methods need formidable recourses for mathematical programming and they are based on heuristics like kernel functions. SFRF, contrary to methods that use dimensionality reduction algorithms as PCA, KPCA, Singular Value Decomposition [26], uses the full dimensional feature vectors. Dimensionality reduction is undesirable, because it might corrupt the information captured in descriptors. The latter is especially important for 3D objects, because geometric information is the only kind of data that we can obtain from the models. The same argument is valid also for methods that use MDA or LDA, with the hope of making objects more discernible in a lower-dimensional feature space. The classification procedure is always more accurate in the original dimensionality if descriptors have discriminative power. Also, methods that assume a predefined probability distribution of the models (e.g. [8]) are inherently based on an assumption that might not be true. These methods advance to simplifications and assumptions that are led by the predefined distribution. The proposed method does not make any such a priory assumptions, but tries to form the feature space based on users' critic. The most relevant method to the SFRF is the feature space warping method [2] (S.W.). The main advantage of the proposed method against [2] is the speed and the accumulation of all users' feedback. S.W. uses the relevance scores of each distinct query session in order to move the objects. After a query session, the information about relevant scores of previous users and query sessions is lost. In this way, users that feed the system with out-of-sense scores have a catastrophic influence on the clustering. On the contrary, the proposed method displaces every object according to the mean of scores that it has been received. That is to say, SFRF depends on the general opinion of all database's users and thus the influence of spurious malevolent or ignorant feedbacks is ameliorated. Furthermore, the proposed method displaces only a small percentage of the objects in the database after each iteration and not all objects as is done in [2]. This enhances significantly the computational speed. PFRF has most of the advantages of SFRF plus it is a purely content-free method. The "descriptors" in this case, are the votes that each object has received from all users, during all RF iterations. Exactly those descriptors are manipulated by the calculated semantic forces and 3D objects are properly displaced in this formed semantic space.

3 The SFRF proposed method

3.1 Conceptual representation of a 3D object

Let us assume that each 3D object, $\mathcal{O}_i$, $i \in [1, N]$, where N is the number of the 3D objects in the database, is represented by two vectors. The first vector is the *feature vector*, $\mathbf{d}_i$, which represents the geometric descriptors that have been extracted from the object. In the current implementation, the magnitude, $\|a_{lm}\|$, of the Spherical Harmonic coefficients [15], $f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm} Y_l^m(\theta, \phi)$ are used to form $\mathbf{d}_i$, where $f(\theta, \phi)$ is a spherical function that represents the 3D model and is decomposed as the sum of its harmonics $Y_l^m(\theta, \phi)$, where $Y_l^m(\theta, \phi) = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_l^m(cos\theta)e^{im\phi}$. The second vector is called the *charging vector*, $\mathbf{q}_i$ and represents the users' relevant judgement for inter-object similarity in the database. Therefore, a conceptual representation of an object is formed as $\mathcal{O}_i = \{\mathbf{d}_i, \mathbf{q}_i\}$.

The charging vector, $\mathbf{q}_i$, for each object, $\mathcal{O}_i$, is formed as follows: For each object $\mathcal{O}_i$, $i \in [1, N]$, where N is the number of the 3D objects in the database, a vector, $\mathbf{q}_i$, is kept. Each cell, $q_i[j]$, of this vector represents the similarity of object $\mathcal{O}_i$ with object $\mathcal{O}_j$ in terms of users' votes. For example, let us assume that a user queries the system with $\mathcal{O}_1$ and the system retrieves a list of, say 7, objects. Let us, also, assume that the user judges as highly relevant the objects $\mathcal{O}_3$ and $\mathcal{O}_4$ with scores "2", as relevant the objects $\mathcal{O}_2$ and $\mathcal{O}_5$ with scores "1" and as non-relevant the objects $\mathcal{O}_6$ and $\mathcal{O}_7$ with scores "-2". Assuming that $N = 7$, then the charging vector for this query and for one user is: $[x, 1, 2, 2, 1, -2, -2]$, where x denotes the self relevance. In the case that another user queries the database with the same object $\mathcal{O}_1$ and his feedback is "2" and "1" for objects $\mathcal{O}_3$ and $\mathcal{O}_4$, respectively, and "0" for the other objects, then the final result for the charging vector $\mathbf{q}_i$ is the vector sum of the old and the new charging vector. That is to say, the final vector, that is kept in memory each moment for each object, is the sum of the charging vectors of each session till that moment. In the above example, the resulting vector after the second session is: $\mathbf{q}_1 + \mathbf{q}_2 = [x, 1, 2, 2, 1, -2, -2] + [x, 0, 2, 1, 0, 0, 0] = [x, 1, 4, 3, 1, -2, -2]$. The physical meaning of this vector is that object $\mathcal{O}_1$ is quite similar to objects $\mathcal{O}_2$, $\mathcal{O}_5$, very similar to objects $\mathcal{O}_3$, $\mathcal{O}_4$ and quite non-similar to objects $\mathcal{O}_6$, $\mathcal{O}_7$.

Thus every object, $\mathcal{O}_i$, is described by a charging vector $\mathbf{q}_i$, whose each cell, $q_i[j]$, represents the cumulative relevance scores for an object $\mathcal{O}_j$, if we use as query the object $\mathcal{O}_i$ and the users give scores to object $\mathcal{O}_j$. The cumulative score is summed over all user sessions and is stored in the database. The relevance scores of each session are added to the corresponding scores of previous sessions and thus each moment we have the inter-object similarities represented by the vector cells. For objects that have not received user feedback the score added is “0”. The more the users who query the database, the more reliable the estimation of the similarities becomes. The absolute value of each cell, $\|q_i[j]\|$, can be interpreted as the magnitude of the charge of a charged element and the sign of each cell, as the sign of the charge. In the aforementioned example of a database with 7 objects and a charging vector $[x, 1, 4, 3, 1, -2, -2]$ for object $\mathcal{O}_1$, it can be elicited that the object $\mathcal{O}_1$ is positive charged against objects $\mathcal{O}_2$,$\mathcal{O}_3$,$\mathcal{O}_4$,$\mathcal{O}_5$ and negative charged against objects $\mathcal{O}_6$,$\mathcal{O}_7$.

3.2 The geometric distance

The geometric distance, r_g, between two objects, $\mathcal{O}_k$ and $\mathcal{O}_l$, is defined as the Minkowski distance between the feature vectors, $\mathbf{d}_k$ and $\mathbf{d}_l$, of these objects. This metric has been proven to give good results in geometric retrieval methods. Therefore, the geometric distance is formulated as: $r_g(\mathcal{O}_k, \mathcal{O}_l) = \sqrt[n]{\sum_{i=0}^{D-1} \|d_k[i] - d_l[i]\|^n}$, where $n \in \Re^+$ and D is the number of descriptors used. Henceforth, the representation r_g is used in the position of the $r_g(\mathcal{O}_k, \mathcal{O}_l)$, when it is obvious which are the referenced objects. The geometric distance can be transformed to similarity by the equation $simil_g = 1 - \frac{r_g}{\max r_g}$ where $\max r_g$ is the maximum geometrical distance between objects in the database. After the semantic clustering, as it is described in Section 3.3, there is no need of a semantic oriented metric since the burden of depicting semantic similarities/disimiarities into the low-level feature space has been carried by the displacement of the objects due to the semantic forces. Actually after the semantic clustering, the choice of the Minkowski metric is the natural choice for a distance metric, as long as the retrieval procedure takes place in the altered but still low-level, feature space.

3.3 The semantic forces and shaping of the feature space

The general idea behind the proposed scheme is the proper displacement of the objects in the geometric feature space in order to gather semantically similar objects together and separate apart semantically non-similar objects. Let us assume that objects which have received positive RF during users’ query session are moved together and negative examples are moved away from positive ones. Each retrieval and each RF iteration, leads to displacement of certain objects and shapes the feature space. An assumption is made that, conversely to the classical electromagnetic field theory, in the case of RF, two positively charged particles (3D object that have both received positive feedback) are attracted to each other and negatively charged particles (3D objects that have both received negative feedback) are repelled away from each other. In case of objects that have received feedback of different sign, no force is being applied since this means that users have voted inconsistently for each of these objects when the other one was used as query. In this case, the system awaits the votes of other users in order to deduct assumptions about the force. Extremely rarely, does this inconsistency arise due to the mean, common sense nature of the charging vector as it is described later. Let us assume that the absolute value of the $q_i[j]$ cell of the charging vector of object $\mathcal{O}_i$, depicts the magnitude of the charge of object $\mathcal{O}_i$, in relation to object $\mathcal{O}_j$, if object $\mathcal{O}_i$ is used as query, and the absolute value of the $q_j[i]$ cell of the charging vector of object $\mathcal{O}_j$, has the symmetric meaning.

Then the force that is being exerted by object $\mathcal{O}_i$ to object $\mathcal{O}_j$ in the feature space, is defined to be:

$$\mathbf{F}_{i,j} = \begin{cases} k\dfrac{\|q_i[j]\| \cdot \|q_j[i]\|}{r_g^p} \cdot \widehat{\mathbf{e}}_{ij}, & \mathcal{O}_i, \mathcal{O}_j \text{ are negatively charged and thus considered} \\ & \text{non relevant} \\ k\dfrac{\|q_i[j]\| \cdot \|q_j[i]\|}{r_g^p} \cdot \widehat{\mathbf{e}}_{ji}, & \mathcal{O}_i, \mathcal{O}_j \text{ are positevely charged and thus considered} \\ & \text{relevant} \\ \text{no force calculation,} & \text{if } q_j[i], q_i[j] \text{ have different sign} \end{cases} \tag{1}$$

where k is a constant, $\widehat{\mathbf{e}}_{ij}$ and $\widehat{\mathbf{e}}_{ji}$ are the unitary direction vectors in the direction of the line $\overrightarrow{\mathcal{O}_i\mathcal{O}_j}$ and $\overrightarrow{\mathcal{O}_j\mathcal{O}_i}$ respectively and $p \in \Re^+$. The geometric distance, r_g, between two objects is defined as the Euclidian distance between the geometric feature vectors of these objects. In general $q_i[j]$ may be different from $q_j[i]$, since this non-symmetrical treatment of the charges, is believed to capture slight variations in human perception in differences between objects $\mathcal{O}_i$ and $\mathcal{O}_j$, when the first or the second is used as query by different users. The direction of the force can be repulsive ($\widehat{\mathbf{e}}_{ij}$) or attractive ($\widehat{\mathbf{e}}_{ji}$), something that is resolved based on whether users have evaluated as relevant or non-relevant the objects under discussion. In the current application, for each RF iteration, the force magnitude is calculated between the query and each object which has received feedback in the current iteration. If one of the judged objects has a total negative feedback and negative charge has also the query, then this object is repelled away from the query (repulsive force) during the current session. Conversely, if an object has a total positive feedback and positive is the charge for the query too, then a attractive force is being exerted to it by the query. In case of different signs in the total charge of the query and a judged object no force is calculated.

In the typical situation after the user has specified his/her query object, the system turns back a ranked list of the most similar objects according to the similarity metric described in Section 3.2. The user observes this list and gives scores to arbitrarily many 3D objects. The score regime can be either binary or gradually e.g. “excellent match” scores “2”, “good match” scores “1", “indifferent match” scores “0”, “poor match” scores “−1”, “no relation at all” scores “−2”. In binary score regime the user has only a relevant/irrelevant judgement option. Those scores are fed to the system and additively accumulated with previous scores from the same query session or other users’ query sessions as indicated in Section 3.1.

After each iteration the query acts as a point charge that exerts certain attractive or repulsive force to objects that have just received feedback. The magnitude of this Semantic Force depends on the accumulated knowledge about the previous users’ but also on the current user’s discriminative power and selection taste for that query. In case this force is attractive, the respective object is moved towards the query in the direction $\overrightarrow{\mathcal{O}\mathcal{Q}}$, where $\mathcal{Q}$ is the query in the geometric feature space and $\mathcal{O}$ the object under discussion. Otherwise, if the force is repulsive, the object is moved away form the query, in the direction of $\overrightarrow{\mathcal{Q}\mathcal{O}}$.

Until this point, we have been reluctant to indicate blindly that after an RF iteration, the query will exert attractive force to all positive examples and repulsive force to all negative ones. Actually, the case where even if a user marks an object, $\mathcal{O}_i$

as relevant to his query, this object is moved away from the query is possible. This can happen in the case where all previous users that queried the database with the same query, have given negative scores to object $\mathcal{O}_i$. This happens exactly because voting information from all users, across all sessions are accumulatively stored in the charging vector of each object. For example if $q_i[j]$ is "100" up to a point, then this means that $\mathcal{O}_j$ is very similar to $\mathcal{O}_i$ when the latter is used as query, according to many users. If now a certain user give a negative mark, say "−2", to $\mathcal{O}_j$ when he queries the database with $\mathcal{O}_i$ the respective charge will become "98" and will still remain positive. This is a very desired property of the proposed RF scheme as it ameliorates the consequences of sporadic, accidental or malevolent, feedback actions that make no sense.

Another issue that is of great concern is the amount of displacement of an object towards or away from the query according to the Semantic Force. The geometrical distance in the denominator of the force can be arbitrarily large or arbitrarily small, since two objects can be very near (the same in the extreme case) or very far away in the feature space. Furthermore, the numerator can also attain very large, without sense values, since the cells of each charging vector are updated additively. In order to alleviate the latter anomaly, the value of each cell, $q_i[j]$, of the i-th charging vector, is normalized with the number of times the users have voted for the relation of $\mathcal{O}_i$ to $\mathcal{O}_j$. Thus, the numerator of the force is the product of two real number values in the space $[-2, 2]$, if the score regime is in the range $[-2, 2]$, which are the mean scores that objects have received. After the normalization of the numerator terms, the possibility of the force being arbitrarily large remain due to the unbounded r in the denominator. An upper bound of r can easily obtained if we scale the components of the feature vector so as the value of each descriptor is in a bounded space e.g. [0,10] or [0,1]. However, the lower bound of r is unavoidable the zero value, since 2 objects can be arbitrarily near, and thus, leading the semantic force to infinite values. In order to elbow through this difficulty, the magnitude of the force in passed through a sigmoid function so as to obtain a percentage value, γ in the space [0,1]. The object, $\mathcal{O}_i$, to be moved is displaced by an amount that is equal to this percentage of the initial Euclidian distance of the object to the query, $\mathcal{Q}$. That is to say, the update of the judged object's, $\mathcal{O}_i$, feature vector, $\mathbf{d_i}$ is:

$$\mathbf{d}_i^{new} = \begin{cases} \mathbf{d}_i^{old} + \gamma \cdot \|\mathbf{d}_q - \mathbf{d}_i^{old}\| \cdot \widehat{\mathbf{e}}_{ji}, & \mathcal{O}_i, \mathcal{Q} \text{ are both positively charged} \\ \mathbf{d}_i^{old} + \gamma \cdot \|\mathbf{d}_q - \mathbf{d}_i^{old}\| \cdot \widehat{\mathbf{e}}_{ij}, & \mathcal{O}_i, \mathcal{Q} \text{ are both negatively charged} \\ \mathbf{d}_i^{old}, & \mathcal{O}_i, \mathcal{Q} \text{ have charges with different sign} \end{cases} \quad (2)$$

where $\widehat{\mathbf{e}}$ is the unit norm, direction vector of the Semantic Force, $\mathbf{d}_i^{old}$ ($\mathbf{d}_i^{new}$) is the old (new) position of the object $\mathcal{O}_i$ in the feature space and $\mathbf{d}_q$ is the position of the query. The sigmoid function has the form: $\gamma(\|\mathbf{F}\|) = \frac{2a}{1+e^{-b\|\mathbf{F}\|}} - a$, where the parameter a and b determine the "saturation" point and the slope of the function respectively and are estimated experimentally.

4 Pure semantic relevance feedback

4.1 Description

The second proposed algorithm takes into consideration only the charging vector. To the extend of authors' knowledge there has been not devised a similar approach to the RF problem till day and thus the proposed scheme is completely innovative. In this method, the charging vectors of each object form the feature space in which the objects belong. The conceptual representation of a database object becomes $\mathcal{O}_i = \{\mathbf{q}_i\}$. The objects are, also, displaced in this method in the (pure now) semantic space, although a different strategy is adopted in this case. The positive examples are moved nearer to the query as: $\mathbf{q}_i^{new} = \mathbf{q}_i^{old} + \acute{\gamma} \cdot \|\mathbf{q}_q - \mathbf{q}_i^{old}\|$, where $\mathbf{q}$ represents the charging vector -new or old- and $\acute{\gamma}$ is a semantic coefficient, which is calculated as presented in Section 4.2. The above formula alters the charge of a positive example in order to be more similar to the query. The degree of change is determined by the percentage of semantic similarity, $\acute{\gamma}$. The negative examples are treated differently. Their charging vector does not change as a whole, as in the previous case, but rather only one dimension is changed, the one, which indicates the relevancy degree between the query and the object. The value of the charging vector along this dimension is set to a negative value (e.g. -2), as the object under discussion has been judged as non-relevant.

4.2 The semantic coefficient

The coefficient $\acute{\gamma}$ indicates the similarity of two charging vectors and thus, it is illustrative of the semantic similarity of two objects. Typically, the semantics of the entities of a database are understood by means of a probabilistic framework or an ontology. In these cases, the semantic similarity is measured by estimating the distance between two probability distribution functions [4] or respectively by defining a metric on the ontology structure [30]. On the other hand, the proposed method of this paper is based on confronting user votes as values of vector dimensions, a fact that offers a genuine semantic interpretation and also alleviates problems of similarity calculation. Therefore, even simple metrics such as L_1 distance can be used (although L_1 has not been used in further experiments since it showed inferior results in the first steps of the implementation). In the current implementation, three metrics are used: the Tanimoto distance (3), the Dice distance (4) and the Soergel distance (5) [3]. These distances are believed to capture the semantic meaning of the distance between two probabilistic or semantic vectors as it is deduced by the experimental results.

$$S_{i,j}^{Tanimoto} = \frac{\sum_{k=1}^{D} q_i[k]q_j[k]}{\sum_{k=1}^{D} q_i^2[k] + \sum_{k=1}^{D} q_j^2[k] - \sum_{k=1}^{D} q_i[k]q_j[k]} \tag{3}$$

$$S_{i,j}^{Dice} = \frac{2\sum_{k=1}^{D} q_i[k]q_j[k]}{\sum_{k=1}^{D} q_i^2[k] + \sum_{k=1}^{D} q_j^2[k]} \tag{4}$$

$$S_{i,j}^{Soergel} = \frac{\sum_{k=1}^{D} \|q_i[k] - q_j[k]\|}{max(q_i[k], q_j[k])} \tag{5}$$

where in all three formulas, $S_{i,j}$ is the similarity value between two objects $\mathcal{O}_i$ and $\mathcal{O}_j$, D is the number of dimensions in the charging vectors and $q_i[k]$ is the value of the k-th dimension of the charging vector of object $\mathcal{O}_i$. When $S_{i,j}^{Tanimoto}$ and $S_{i,j}^{Dice}$ metrics are used, then $\acute{\gamma} = S_{i,j}^{Tanimoto}$ or $S_{i,j}^{Dice}$. On the contrary, when the $S_{i,j}^{Soergel}$ is used, then $\acute{\gamma} = 1 - S_{i,j}^{Soergel}$ because the Soergel metric indicates distance instead of similarity.

5 Cold start problem and annotation propagation

Today's database users need to store huge amounts of multimedia data, which must be given some kind of annotation, if one desires to re-utilize the deposited knowledge with the maximum effectiveness. The typical annotation process is done manually and demands many precious man-hours. One of the most important and challenging problems in multimedia databases is the propagation of the already stored information in order to annotate new forthcoming objects.

In the RF methods that are presented in this paper, the vector of descriptors of each object, whether it is a pure semantic one as in PSRF or geometric as in SFRF, is permanently changed with every RF iteration. This way the initial position of the database objects is changed to a new updated position and thus semantic clusters are formed in the vector space. The displacement of the initial object position is liable only to the inter-object semantic powers. The formation of semantic clusters leads, also, to the clearer redefinition of the current, human defined classification (e.g. humans, animals, cars etc.). Thus, the annotation procedure in the proposed algorithms is achieved indirectly through the proper displacement of the objects and therefore, it is not a manual task, but rather an automated procedure, which is transparent to the user.

Nevertheless, it is clear from the previous paragraph that if one wants to add new objects to an already formed database, the semantic space of which has already been distorted, cannot just insert the initial descriptor vectors of the new objects. The initial vector of a new object, either the geometric descriptor vector as it is extracted from a certain geometric algorithm or an all "1"s vector if it is a pure semantic vector, must catch up with the displacements and the changes that the other objects' vectors have undergone. If the initial vectors of new object were inserted into the retrieval algorithm, the retrieval accuracy would be rather poor. This problem suggests the cold start problem and can be solved by propagating the already stored knowledge, that is to say the amount of displacement of the already changed objects, to the newcomers—the objects to be added.

In order to alleviate the aforementioned problems, a method for annotation propagation by estimating the proper displacement of a new object, is devised in this paper. Let us assume that the displacement vectors $\delta\mathbf{d}_i^{old}$, of all objects $\mathcal{O}_i$, $i \in [1, N]$ that have already undergone the process of RF and thus their position has been altered, form an $N \times D$ matrix $\delta\mathbf{R}^{old}$:

$$\delta\mathbf{R}^{old} = \begin{pmatrix} \delta d_1(1) & \delta d_1(2) & \dots & \delta d_1(D) \\ \delta d_2(1) & \delta d_2(2) & \dots & \delta d_2(D) \\ \vdots & \vdots & \vdots & \vdots \\ \delta d_N(1) & \delta d_N(2) & \dots & \delta d_N(D) \end{pmatrix} \tag{6}$$

where N is the number of already displaced objects, D is the number of descriptors or equivalently the number of dimensions of the feature space, and $\delta d_i(j) = d_i^{final}(j) - d_i^{initial}(j)$ is the difference of the i-th object's initial feature vector along the j-th dimension minus the i-th object's final (updated) feature vector along the j-th dimension.

The annotation propagation method that is proposed in this paper, is based on the assumption that the displacement vectors, $\delta\mathbf{d}_i$, that are formed by objects, $\mathcal{O}_i$, which belong to the same semantic cluster (e.g. they are all human 3D models), are statistically and spatially correlated to each other. This assumption is quite common sense, if we take into consideration that objects, which belong to the same semantic class are displaced towards similar positions in the feature space by the proposed RF algorithms. This correlated displacement vectors are used to estimate a proper displacement vector for the new objects that are about to be added in the database.

The estimation of the displacement vector, $\delta\mathbf{d}_i^{new}$, for a new object, $\mathcal{O}_i^{new}$, with $i > N$, is achieved by linearly predicting the best displacement (in a mean square error sense) along each dimension. For example, in (6), the values $\delta d_1(1), \delta d_2(1), \ldots, \delta d_N(1)$ form a series, which can be confronted as a kind of time-series. In order to be consistent with the linear prediction theory, an ordering is imposed on the objects that are to form the series which are to be used to predict the optimum displacement of the newly-added object. The objects that are already in the database are ordered in a maximum displacement-first manner. Thus, the object which has the maximum total displacement, $\sqrt{d^2(1) + d^2(2) + \ldots + d^2(D)}$, is the $i-1$ predecessor of the new object. The immediately older $(i-2)$ predecessor has a smaller total displacement, etc. This ad-hoc ordering is adopted in order to give larger importance to objects that have been largely displaced, since this means that they have been manipulated in a greater degree by the semantic forces and thus they enclose a greater amount of user knowledge than other objects. In the following paragraph the linear prediction method that is used in the current implementation of the suggested RF algorithms is described. Linear prediction has many applications, with the most renounced the linear prediction coding [23]. Similar series to the aforementioned one, are formed for each dimension. Therefore, for each dimension j, a series $\delta d_1(j), \delta d_2(j), \ldots, \delta d_N(j)$ can be thought. In the following equations, each of the D formed series is treated in the same manner so for simplicity's sake we denote the formed series as $\{y_n\}$, where $y_i = \delta d_i^{old}(j)$ for $i < N$ and $y_i = \delta d_i^{new}(j)$ for $i > N$, that is to say for the new objects that are about to be added.

The desired displacement, y_n, of the forthcoming object $\mathcal{O}_n^{new}$, along the j-th component can be estimated from the previous M sample values (the displacement values of the M previously already displaced objects) as: $y_n = \sum_{i=1}^{M} a_i y_{n-i}$, where M is the order of the predictor. a_i are the filter coefficients, which must be selected in such a way so that the mean square error $r^2 = \left(y_n - y_n^{original}\right)^2$ is minimized. Equalizing the derivatives of the expected r^2, with respect to the coefficients a to 0, M equations are derived: $\frac{\partial}{\partial a_k}\mathrm{E}\left\{(y_n - \sum_{i=1}^{M} a_i y_{n-i})^2\right\} = 0$. From the partial derivatives above, the following equations are deduced:

$$\sum_{i=1}^{M} a_i \mathrm{E}\{y_{n-i} y_{n-k}\} = \mathrm{E}\{y_n y_{n-k}\} \tag{7}$$

Let us assume, now, that the $\{y_n\}$ sequence is stationary. Taking that into consideration, we have:

$$\mathrm{E}\{y_{n-i} y_{n-k}\} = R_{yy}[\|i-k\|] \tag{8}$$

where $R_{yy}[\ldots]$ is the correlation function. If $\mathbf{R_{top}}$, is the Toeplitz matrix of the first M samples of the $\left\{R_{yy}[\|i-j\|]\right\}$, then the coefficients $a_i, i \in [0, M-1]$, can be estimated from the equation:

$$\mathbf{R_{top}A} = \mathbf{P} \tag{9}$$

where:

$$\mathbf{R_{top}} = \begin{pmatrix} R_{yy}[0] & R_{yy}[1] & R_{yy}[2] & \ldots & R_{yy}[M-1] \\ R_{yy}[1] & R_{yy}[0] & R_{yy}[1] & \ldots & R_{yy}[M-2] \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ R_{yy}[M-1] & R_{yy}[M-2] & R_{yy}[M-3] & \ldots & R_{yy}[0] \end{pmatrix} \tag{10}$$

$$\mathbf{A} = \begin{pmatrix} a_1 \\ \vdots \\ a_M \end{pmatrix} \tag{11}$$

$$\mathbf{P} = \begin{pmatrix} R_{yy}[1] \\ \vdots \\ R_{yy}[M] \end{pmatrix} \tag{12}$$

In order to solve the aforementioned equations in the current implementation, the efficient Levinson-Durbin [18, 23] recursive algorithm is used.

However, the method for annotation propagation that has just been denoted assumes that the class of the object that is about to be added in the database is known. Thus, for example, in the case that a user knows that s/he is going to insert a 3D model of a human, s/he submits this information along with the model into the system. Then a prediction for the proper displacement of the newly added model is estimated, based on the displacements of the other 3D human models that already exist in the database. Although, in the case where a user does not know or does not want to provide the new object's category to the system, there must be a way to determine the predicted displacement vector to be followed by the newly added object. In this paper, this problem is solved with an initial classification of the newly added object. The object is categorized into one of the existing semantic classes and a displacement vector is estimated from the population of this class. If the classification is correct, then the position of the new object is properly updated and the object catches up with the rest of its "family". In case of a misclassification, the new object follows the predicted displacement vector of a foreign class and thus it is being relocated amongst non-relevant objects in the feature space. This latter case does not pose a considerable problem in the proposed RF framework, because an "alien" object among object of the same class will be easily repelled away and attracted by relevant objects. Therefore, the user himself can either define the category of the newly added object, or rely on the classification scheme. In the latter case and if the classification happens to be erroneous, further RF iterations are needed in order to achieve the best results.

The classification method that is used in the current implementation, is a modification of the Reduced Coulomb Energy (RCE) network [26]. RCE is a non-parametric classification method. Its advantage, in the RF problem, stems from the fact that RCE confronts the classification problem in a way, which is intermediate between Parzen-window method and the k-nearest neighbor, the two main members of the nonparametric family of methods. Whereas Parzen-window method uses a fixed window throughout the feature space, which could lead to difficulties of how to define the best size of the window and the k-nearest neighbor method addresses this problem by adjusting the region based on the density of the points and is largely dependent on the distance metric that will be used, the RCE method tries to adjust the size of the window during training according to the distance to the nearest point of a different category. The modified, based on RCE, algorithm that is implemented in this paper, is called Sphere Covering Pattern Classification and is presented in [20, 32].

6 Experimental results

In order to evaluate the proposed method in terms of retrieval accuracy, the ITI Database [9] and the Princeton Shape Benchmark (PSB) [25] were used. The ITI database consists of 544 3D objects which are semantically clustered in 13 categories. From these categories the first 12 include objects that are semantically similar in each class, whereas the 13th category consists of semantically non-similar objects. This 13th category behaves actually as noise in the whole process. The PSB consists of 1814 3D objects which are categorized in 4 tiers e.g. animals/4-legs/mammals/dogs. Each sub-category consists of at least 4 objects and maximum 100 objects.

The relevant scores are in the range $[-2, 2]$ for all experiments and the SH curve is the Precision-Recall curve for the retrieval process with the original geometrical spherical harmonic descriptors [15], using L_1 as distance metric. Also, in current implementation the values $p = 1$, $p = 0.5$ and $p = 2$ in (1), have been experimentally selected. The platform used for the experiments was a Dual Core Duo 1.83 GHz/1 GB RAM/WinXP machine running Visual Studio 2005.

6.1 Evaluation of SFRF

The proposed SFRF algorithm has been tested in several ways. In the first test our goal is to determine the parameters of the sigmoid function and the range of the relevant scores. Each object acts as query for one time. The objects are picked randomly until every 3D model has been used as query for one time. Each query is used for It=10 RF iterations. During every RF iteration the first T=20 objects take relevant scores from the user. In Fig. 1a are depicted the Precision-Recall (PR) diagrams for different values of the b parameter of the sigmoid function. In Fig. 1b are given the corresponding data for the PSB. The value $b = 0.002$ is considered the best overall in the ITI database and is kept for further experiments. The value $b = 0.025$ is considered the best for the PSB. The value a, of the sigmoid function is kept constant, $a = 1$, through all experiments.

Further experiments have been conducted in order to evaluate the proposed algorithm in terms of the number of RF iterations (It). In order to test the aforementioned property of the proposed algorithm, we compare our method to the feature space warping method [2] (S.W.). This method was chosen for comparison, because it is the only method that changes the initial descriptors of the 3D objects and keeps

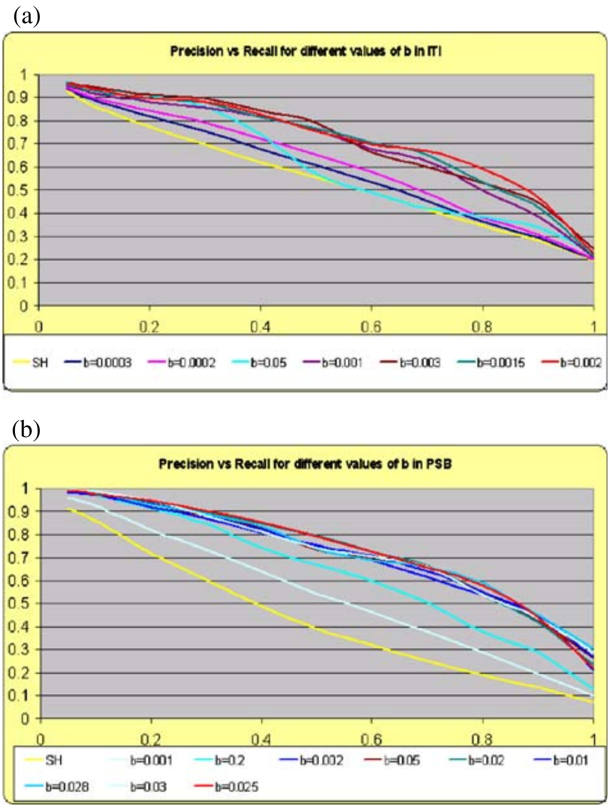


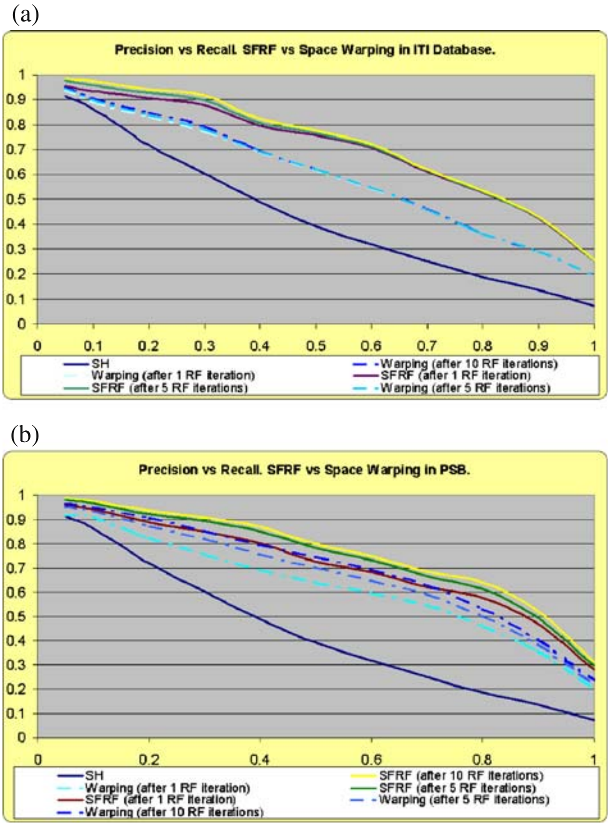
Fig. 1 Precision-Recall curves for various values of parameter b of the sigmoid function. Parameter a, is kept in value $a = 1.0$. **a** ITI database, **b** PSB

the changed ones (as the proposed method also does). The best parameters for the warping method were experimentally found to be $\gamma = 0.02$, $c = 20\pi$ ($\gamma = 0.009$, $c = 6\pi$) for the ITI database (PSB, respectively). The experimental results show that the proposed method is superior to its counterpart in both databases.

In Fig. 2, the proposed method is compared to the S.W. and to Spherical Harmonics geometric method, in terms of various values of the parameter It. The results are extracted for the ITI database and for the PSB after running the SFRF for those datasets and giving feedback in the first 20 retrieved objects. Each object acts one time as query. It is obvious that the most noticeable improvement in the PR diagrams is achieved in the first RF iteration. This is a very desirable property since our purpose is to semantically cluster the database with the least possible number of steps. In a database system that receives queries from a huge amount of users, it is certain that the performance becomes near perfect after a short time. Also, in all PR curves the proposed method significantly outperforms the feature space warping [2].

After the aforementioned experiments, the performance of the SFRF in terms of the number of times that each database object acts as query, is investigated. In order to evaluate the SFRF response the following experiment is performed: Every object acts as query for K=2, K=5 and K=10 times. The SFRF method and the S.W. method run for It=1 RF iteration, that is to say, when an object is selected to act as query only one time, a ranked list is retrieved and feedback is given to the

Fig. 2 Precision-Recall curves for various number of SFRF iterations compared to the feature space warping method. Results are drawn for the **a** ITI database and the **b** PSB



T=20 top ranked objects. Results are illustrated in Fig. 3 and show that the retrieval accuracy is dramatically enhanced. This gives rise to the argument of the authors that as more and more users "feed" the system with their scores and relevance preferences, the retrieval accuracy of the proposed technique is getting better. In a database system that receives queries from a huge amount of users, it is certain that the performance becomes near perfect after a short time. In the same figure is proved that the proposed method outperforms the feature space warping in these experiments.

Lastly, one more experiment is performed in order to evaluate the performance of the SFRF algorithm in terms of different values of the parameter T, that is to say the number of objects that receive feedback in each RF iteration. Firstly, using as dataset the ITI database, every object acts as query for $K = 1$ time. The SFRF method is used for It=1, RF iteration and feedback is given to the $T = 10$, $T = 20$, $T = 35$ and $T = 50$ top ranked objects. Results are illustrated in Fig. 4a where for $T = 50$, the results are quite extraordinary. This is possible because in ITI database most categories have under 50 objects each one. Because of this restriction and the huge amount of feedback (with judging 50 objects one actually does not avoid the manual annotation of the database) the results are nearly perfect. It is obvious that the more the number of objects that receive feedback grows, the more accurate the

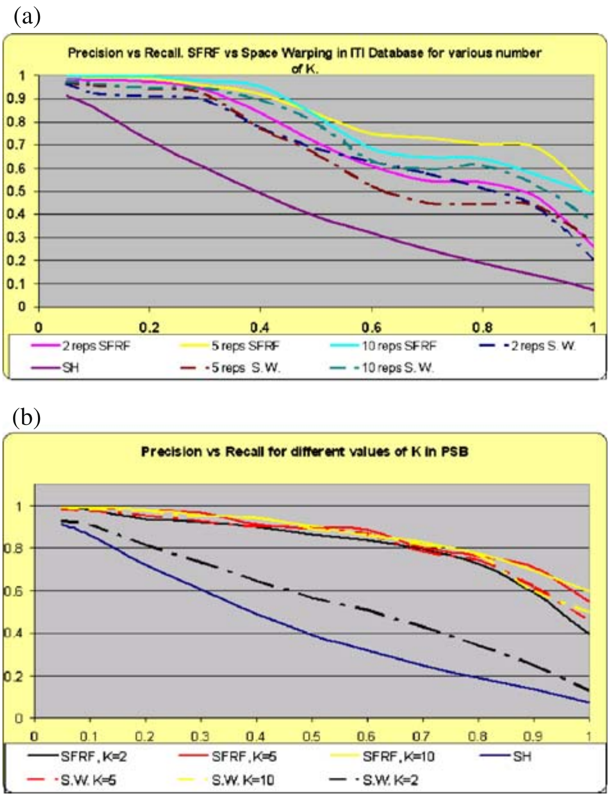


Fig. 3 Precision-Recall curves for various values of the parameter K, using SFRF and Space Warping. Results are drawn for the **a** ITI database and the **b** PSB

retrieval process becomes. Although it is cumbersome to manually provide feedback to large quantities of objects, it is certain that in the long run and as the users peruse the system and give feedback to various objects, the process of giving feedback to large numbers of objects is being "simulated".

Furthermore, the SFRF method and the S.W. method run for It=1, RF iteration and feedback is given to the $T = 10$, $T = 15$, $T = 20$ top ranked objects. Those numbers are closer to reality, because users tend to annotate only few samples. The results indicate the superiority of the proposed algorithm. Lastly, SFRF ran for one iteration for every model in order to calculate the mean timings. The retrieval and RF steps took 233.1 and 0.04 ms respectively.

In order to evaluate the classification accuracy of the proposed scheme, the hierarchical clustering [26], an unsupervised learning algorithm, was used to estimate the number of natural classes in the ITI database, before and after the RF procedure has been applied. Initially 154 classes were found (cophenetic coefficient 0.884), while after running SFRF with $T = 20$, $K = 10$, It=1, 36 classes were found (cophenetic coefficient 0.879). The reduction of the number of natural formed clusters along with the improved precision-recall results, indicate that the proposed algorithm performs a very successful semantic clustering.

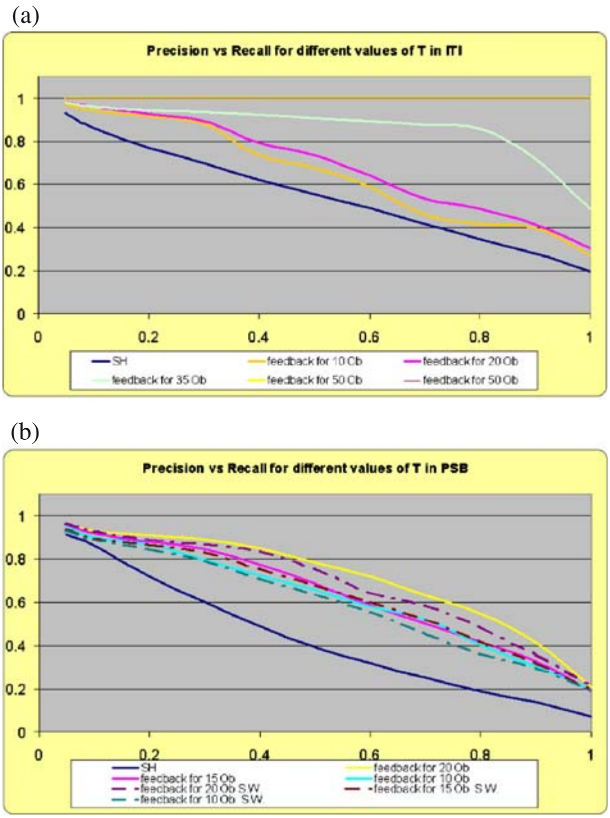


Fig. 4 Precision-Recall curves for various values of the parameter T, using SFRF and Space Warping. Results are drawn for the **a** ITI database and the **b** PSB

6.2 Evaluation of PSRF

The second proposed algorithm of this paper, the PSRF, works in a purely semantic feature space, which is constructed by each objects' semantic vector. The components of these vectors represent user votes, which indicate the inter-object semantic relationships. Those vectors can be seen as the position of the corresponding object in the semantic feature space. As has, already, been explained in Section 4, the objects change position according to (4.1) and thus, similar objects attain faster the appropriate votes, without user intervention. The factor $\acute{\gamma}$ in (4.1) depends on the distance metric that is used. Similarly, the accuracy of the final retrieval depends on the kind of objects that are "near" the query; however, "near" depends on the distance metric, too. In current implementation the metrics that has been mentioned in Section 4.2, are used both for the estimation of the parameter $\acute{\gamma}$ and for the final retrieval.

In order to find the best semantic metric, we experiment with all the nine combinations of the selected distance measures for retrieval and parameter estimation. Those combinations for the retrieval and estimation are the tuples (Soergel, Dice), (Soergel, Tanimoto), (Soergel, Soergel), (Tanimoto, Tanimoto), (Tanimoto, Dice), (Tanimoto, Soergel), (Dice, Dice), (Dice, Tanimoto), (Dice, Soergel). The ITI database is used as dataset since it has the advantage against the PSB that it contains a

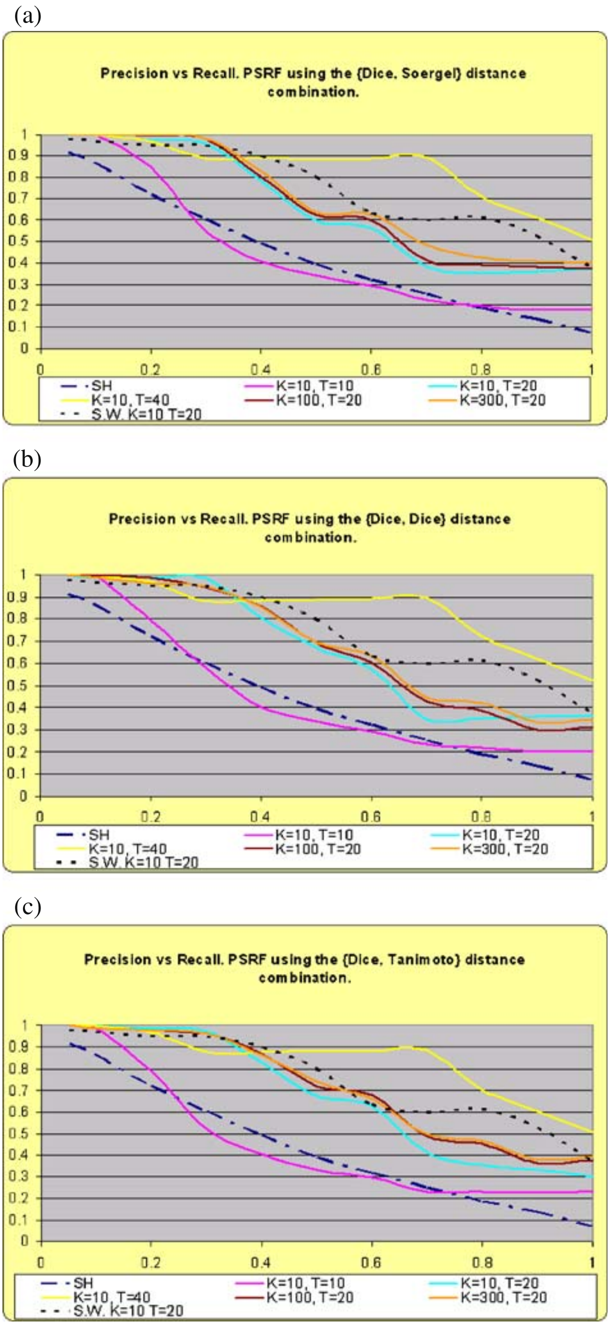
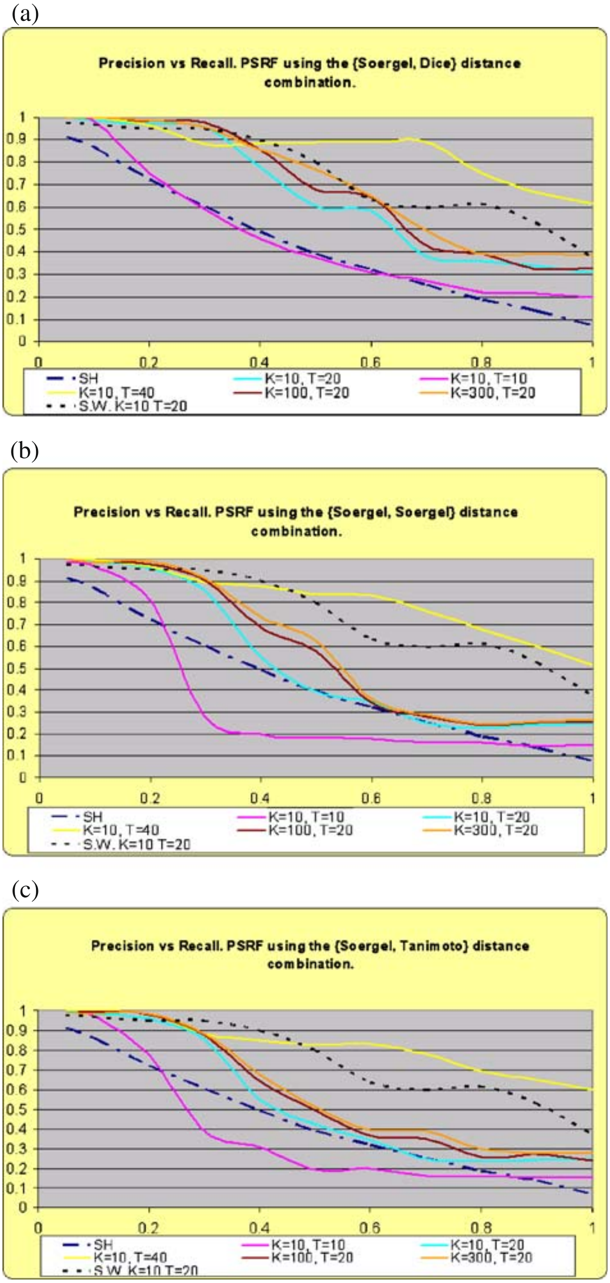


Fig. 5 Precision-Recall results for various values of the parameter K for the ITI database. Results are drawn for the PSRF and for the combinations (Dice, Soergel) (**a**), (Dice, Tanimoto) (**c**), (Dice, Dice) (**b**) of the retrieval and the prediction distance measure

category with various objects, which acts as "noise" in the RF process. This category-noise is ideal to test a pure semantic algorithm, such as the proposed one. For every metric combination the evaluation method is as follows: Firstly, every object in the

Fig. 6 Precision-Recall results for various values of the parameter K for the ITI database. Results are drawn for the PSRF and for the combinations (Soergel, Dice) (**a**), (Soergel, Soergel) (**b**), (Soergel, Tanimoto) (**c**) of the retrieval and the prediction distance measure



database acts as query for $K = 10$ times and feedback is given to the $T = 10$, $T = 20$, $T = 40$ top ranked objects. Afterwards, every object in the database acts as query for $K = 10$, $K = 100$, $K = 300$ times and feedback is given to the $T = 20$ top ranked objects. This way, the PSRF is evaluated in terms of the number, T, of the objects

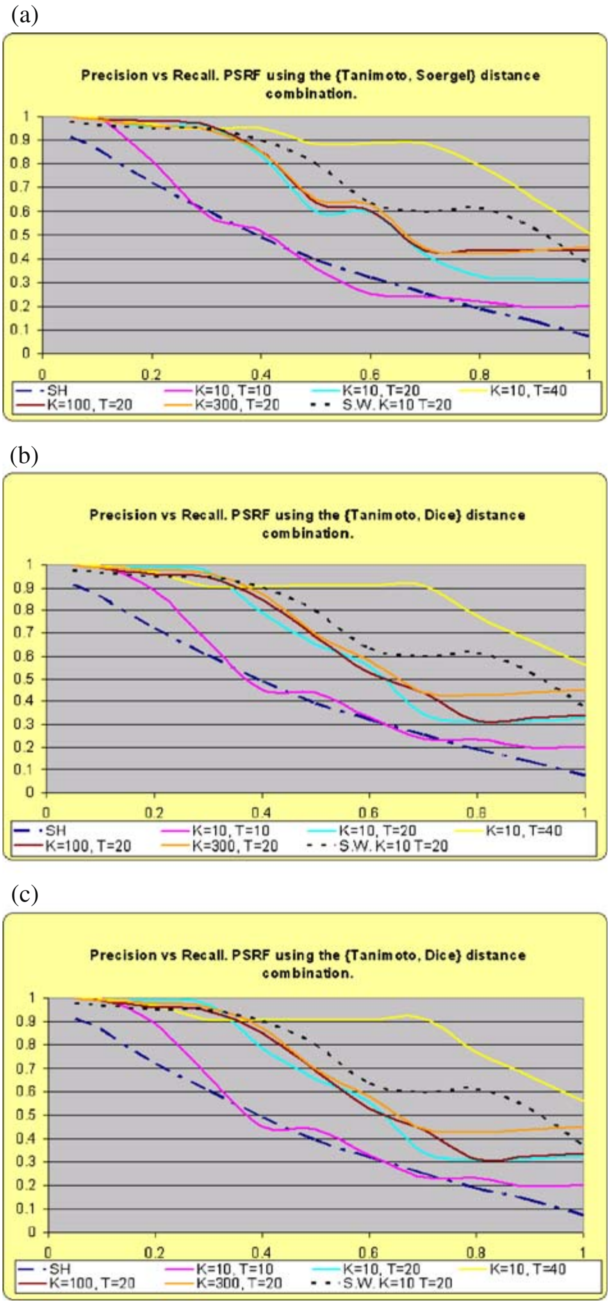


Fig. 7 Precision-Recall results for various values of the parameter K for the ITI database. Results are drawn for the PSRF and for the combinations (Tanimoto, Dice), (Tanimoto, Tanimoto), (Tanimoto, Soergel) of the retrieval and the prediction distance measure (**a**, **b**, **c**)

that each user marks during an RF step and in terms of the number of times an objects has been used as query. The outcome of the PSRF is compared against pure-geometric retrieval (SH) and against the space warping RF method. The results of the Precision-Recall evaluation are illustrated in Figs. 5, 6 and 7. The best overall

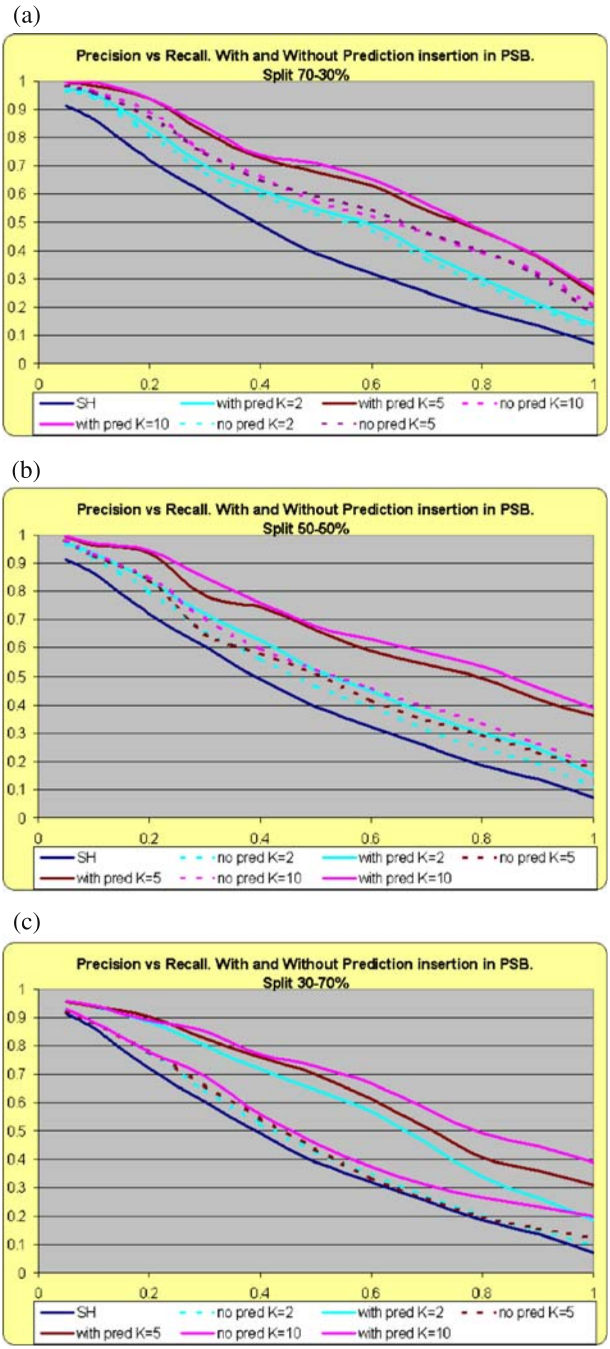


Fig. 8 Precision-Recall curves for various values of the parameter K, after having given feedback to the **a** 70%, **b** 50% and **c** 30% of objects and added the rest with and without prediction

performance have been marked by the (Tanimoto, Tanimoto) combination. The Precision-Recall curves for the experiments with this combination are illustrated in Fig. 7c. In this figure is depicted the fact that PSRF needs either more RF iterations or more objects to get relevance scores during each iteration. For example, for $K = 10$ and $T = 10$, the performance of the RF algorithm is not clearly better that the pure geometric Spherical Harmonics algorithm. However, as the number, K, of the times that each object is used as query grows, or as the number, T, of the objects, which get annotated grows, the performance is boosted. In the case of $T = 40$, the performance is near perfect. As it has been denoted in Section 6.1, difficult and unlike, though it may be for a user to give feedback to 40 objects, this kind of behavior can be approximated in a real situation, due to the fact that users tend to navigate through a wide range of retrieved objects and to give feedback to objects that are not close to the retrieved ranked list. The S.W. RF algorithms for $K = 20$, $T = 20$ outperforms the proposed PSRF algorithm for $K = 20$, $T = 20$. However, the proposed pure semantic algorithm offers invaluable advantages (e.g. there is no need for feature extraction) against a geometric based RF framework as the S.W. and also it illustrates excellent performance as soon as a large number of users offer their opinion for the inter-object relevancies (e.g. PSRF for $T = 40$ greatly outperforms the S.W.). Lastly, PSRF ran for one iteration for every model in order to calculate the mean timings. The retrieval and RF steps took 123.4 and 0.0078 ms respectively.

6.3 Evaluation of the proposed annotation propagation scheme

In order to evaluate the proposed annotation propagation scheme that is described in Section 5, the PSB is used as dataset. Firstly, the 70% of objects of each class, are randomly kept in the database. Every one of those object is used as query. All objects are used as query for 2, 5 and 10 times and during each iteration step the user gives feedback to the 10 most relevant retrieved objects. After each stage (after each object has been used 2 times as query, after 3 times etc.) the other 30% of objects are added to the database. Then, the retrieval accuracy in the whole database is measured in terms of Precision-Recall curves. The experiment is repeated for a 50–50% split and for a 30–70% split. In Fig. 8a, b and c are illustrated the Presicion-Recall curves for the aforementioned experiments. It is obvious from the presented results that annotation propagation is achieved through the prediction of the optimal displacement for a new object. Precision-Recall curves of the database after the insertion of the new objects are certainly worse than the corresponding curves after the insertion of the new objects in positions determined by the predicted optimum displacement. Also, the figures are indicative of the fact that the more the objects are displaced in the database the more necessary is a proper displacement of the new objects. The latter is deduced from the fact that the difference in the PR curves is larger for $K = 5$ or $K = 10$ than for $K = 2$.

7 Conclusion

We have proposed two methods for implementing RF by interpreting each object as a charged element positive or negative dependent on the users' critic. The charged elements apply forces to each other in a way that semantically clusters are formed and the retrieval quality is enhanced. The performance of the first algorithm depends largely on the parameters of the sigmoid function which are found experimentally, whereas the performance of the second, context free method, depends mainly on the number of objects that are given feedback to, by the users, in each iteration. It depends, also, on the number of times that each object is used as query. The proposed method, SFRF, outperforms the Bang and Chen feature space warping [2] as shown by the experimental results and the PSFR, the second innovative method, illustrated very good performance on the performed experiments.

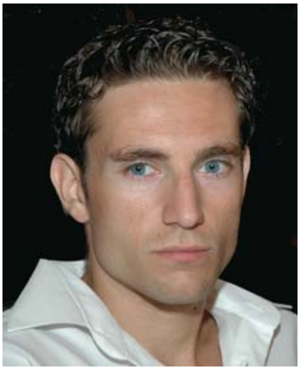
PSFR, being a purely semantic method in the sense that it only utilizes user's cognitive suggestions, belongs to a wider family of research approaches which try to explore human cognition. The semantic gap which is formed due to the lack of our understanding about the function of human brain, can be filled up to a degree by our proposed RF method, since the semantic accuracy of the retrieved results is greatly enhanced, as it is indicated by the experimental results. Therefore, users can interact with 3D databases in a semantic manner, only by giving an example of what they desire to retrieve. The proposed system is able to retrieve the most semantically similar objects.

Moreover, a novel annotation propagation scheme was introduced in this paper. This algorithm is based on a linear prediction method, which estimates the best (in a mean square sense) displacement vector for a newly added object into the database. The estimation of the proper displacement is achieved by taking into consideration the previous displacement vectors of the objects that are already in the database.The experiments that have been conducted in order to evaluate the proposed annotation propagation scheme have indicated that it properly captures the semantic meaning of non-annotated objects and places them near to the already formed semantic clusters.

References

1. Atmosukarto I, Leow WK, Huang Z (2005) Feature combination and relevance feedback for 3d model retrieval. In: Proc. of the 11th int. conf. multimedia modelling. IEEE Computer Society, Washington, DC, pp 334–339
2. Bang HY, Chen T (2002) Feature space warping: an approach to relevance feedback. In: IEEE int. conf. on image processing. Rochester, 22–25 September 2002, pp 968–971
3. Barnard JM, Willett P, Downs GM (1998) Chemical similarity searching. J Chem Inf Comput Sci 38(6):983–996
4. Blachnik M, Duch W, Wieczorek T (2005) Probabilistic distance measures for prototype-based rules. In: Proc. of the 12th int. conf. on neural information processing. Taipei University Press, Taiwan, 30 October–2 November 2005, pp 445–450
5. Chan LW, Chiang CC, Hung YP (2005) Content-based object movie retrieval by use of relevance feedback. In: Proc. int. conf. on image and video retrieval. Singapore, July 2005, pp 425–434
6. Chen SC, Hunag X, Shyu ML (2003) Incorporating real-valued multiple instance learning into relevance feedback for image retrieval. IEEE Int Conf Multimed Expo 2:321–324
7. Chen T, Liu D (2005) Probabilistic relevance feedback with binary semantic feature vectors. In: IEEE proc. int. conf. acoustics, speech and signal processing. Philadelphia, 18–23 March 2005
8. Cox IJ, Miller ML, Omohundro SM, Yianilos PN (1996) Pichunter: Bayesian relevance feedback for image retrieval. In: Proc. of the int. conf. on pattern recognition. IEEE Computer Society, Washington, DC, p 361
9. Daras P, Zarpalas D, Tzovaras D, Strintzis MG (2006) Efficient 3-d model search and retrieval using generalized 3-d radon transforms. IEEE Trans Multimedia 8(1):101–114
10. Flickner M, Sawhney H, Niblack W, Ashley J, Huang Q, Dom B, Gorkani M, Hafner J, Lee D, Petkovic D, Steele D, Yanker P (1995) Query by image and video content: the qbic system. Computer 28(9):23–32
11. He X, Ma WY, King O, Li M, Zhang H (2002) Learning and inferring a semantic space from user's relevance feedback for image retrieval. In: Proc. of the tenth ACM int. conf. on multimedia. Juan les Pins, 1–6 December 2002, pp 343–346

12. Huang TS, Chen Y, Zhou X (2001) One-class svm for learning in image retrieval. In: IEEE int. conf. on image processing. IEEE, Piscataway, pp 34–37
13. Ishikawa Y, Subramanya R, Faloutsos C (1998) MindReader: querying databases through multiple examples. In: Proc. 24th int. conf. very large data bases. New York, 24–27 August 1998, pp 218–227
14. Kanade T, Uchihashi S (2004) User-powered content-free approach to image retrieval. In: Proc. of int. symposium on digital libraries and knowledge communities in networked information society. Tsukuba, March, pp 24–32
15. Kazhdan M, Funkhouser T, Rusinkiewicz S (2003) Rotation invariant spherical harmonic representation of 3d shape descriptors. In: Proc. of the 2003 eurographics/ACM SIGGRAPH symposium on geometry processing. Eurographics Association. Aire-la-Ville, pp 156–164
16. Leifman G, Meir R, Tal A (2005) Semantic-oriented 3d shape retrieval using relevance feedback. Vis Comput 21(8–10):865–875
17. Liu D, Chen T (2006) Content-free image retrieval using bayesian product rule. In: IEEE int. conf. on multimedia and expo. IEEE, Piscataway, pp 89–92
18. Ljung L (1987) System identification: theory for the user. Prentice-Hall, New York
19. Milanfar P, Farsiu TH (2007) Kernel regression for image processing and reconstruction. IEEE Trans Image Process 16(2):349–366
20. Neskovic P, Wang J, Cooper LN (2007) Bayes classification based on minimum bounding spheres. Neurocomput 70(4–6):801–808
21. Oh S, Gyo Chung M, Sull S (2004) Relevance feedback reinforced with semantics accumulation. In: Int. conf. on image and video retrieval. Dublin, 21–23 July 2004, pp 448–454
22. Rui Y, Huang TS, Mehrotra S (1998) Relevance feedback techniques in interactive content-based image retrieval. In: Storage and retrieval for image and video databases. San Jose, 28–30 January 1998, pp 25–36
23. Sayood K (2003) Introduction to data compression. Morgan Kaufman, San Mateo
24. Schölkopf B, Smola A, Müller KR (1998) Nonlinear component analysis as a kernel eigenvalue problem. In: Neural computation, vol. 10. MIT, Cambridge, pp 1299–1319
25. Shilane P, Min P, Kazhdan MM, Funkhouser TA (2004) The princeton shape benchmark. In: Int. conf. on shape modeling and applications. IEEE Computer Society, Piscataway, pp 167–178
26. Stork D, Duda R, Hart P (2000) Pattern classification, 2nd edn. Wiley-Interscience, New York
27. Su Z, Zhang HJ, Li S, Ma S (2003) Relevance feedback in content-based image retrieval: Bayesian framework, feature subspaces, and progressive learning. IEEE Trans Image Process 12(8):924–937
28. Tian Q, Hong P, Huang TS (2000) Update relevant image weights for content-based image retrieval using support vector machines. In: Int. conf. on multimedia and expo. IEEE, Piscataway, pp 1199–1202
29. Tong S, Chang E (2001) Support vector machine active learning for image retrieval. In: ACM Multimedia. ACM, New York, pp 107–118
30. Tsang V, Stevenson S (2004) Calculating semantic distance between word sense probability distributions. In: Proc. of conf. on computational natural language learning. Boston, May 2004, pp 81–88
31. Vapnik VN (1998) Statistical learning theory. Wiley-Interscience, New York
32. Wang J, Neskovic P, Cooper LN (2006) A minimum sphere covering approach to pattern classification. In: Int. conf. of pattern recognition. Hong Kong, 20–24 August 2006, pp 433–436
33. Zitnick C (2003) Computing conditional probabilities in large domains by maximizing Renyi's quadratic entropy. Ph.D. thesis, Robotics Institute, Carnegie Mellon University, May
34. Zhang C, Chen T (2002) An active learning framework for content-based information retrieval. IEEE Trans Multimedia 4(2):260–268
35. Zhang HJ, Su Z (2001) Improving cbir by semantic propagation and cross modality query expansion. In: Proc. of the int. workshop on multimedia content-based indexing and retrieval. Brescia, 19–21 September 2001, pp 83–86



Efstathios Onasoglou received the Diploma degree in Electrical and Computer engineering from the Aristotle University of Thessaloniki, Greece, in 2006. He then joined the Informatrics and Telematics Institute in Thessaloniki, where he worked as an Associate Researcher for the year 2006–2007. He is currently pursuing an MSc in Communications & Signal Processing at Imperial College London. His main research interests include 3D signal processing, content-based/content-free search and retrieval. He is a Member of the Technical Chamber of Greece.

Petros Daras was born in Athens, Greece, in 1974. He is a Researcher Grade D' at the Informatics and Telematics Institute. He received the Diploma degree in electrical and computer engineering, the M.S. degree in medical informatics, and the Ph.D. degree in electrical and computer engineering from the Aristotle University of Thessaloniki, Greece, in 1999, 2002, and 2005, respectively. His main research interests include computer vision, search and retrieval of 3D objects, the MPEG-4 standard, peer-to-peer technologies, and medical informatics. He has been involved in more than 10 European and national research projects. He is a Member of the Technical Chamber of Greece.